

РАЗРАБОТКА МЕТОДИКИ ЭКСПЕРТНОЙ ОЦЕНКИ ЦИФРОВОГО ИНДЕКСА ТРОЛЛИНГА В СОЦИАЛЬНЫХ СЕТЯХ

Мелешкина И.И., Бегичева С.В.

ФГБОУ ВО «Уральский государственный экономический университет»,
Екатеринбург, Россия

Аннотация. В статье рассматривается такая форма современной виртуальной агрессии, как троллинг. Описаны основные характеристики данного явления, способы распространения и влияния данной технологии в России. Приведены примеры методов борьбы с кибер-агрессией в социальном пространстве, а также, описана разработка методики выявления троллинговых аккаунтов в социальных сетях и проведена апробация в Интернет-пространстве.

Ключевые слова: цифровая экономика, цифровизация, виртуальный мир, виртуальная агрессия, кибер-агрессия, троллинг, цифровой индекс троллинга.

DEVELOPMENT OF A METHODOLOGY FOR EXPERT ASSESSMENT OF THE DIGITAL TROLLING INDEX IN SOCIAL NETWORKS

Meleshkina I.I., Begicheva S.V.

Ural State Economic University, Yekaterinburg, Russia

Abstract: The article examines such a form of modern virtual aggression as trolling. The main characteristics of this phenomenon, the methods of distribution and influence of this technology in Russia are described. Examples of methods of combating cyber-aggression in social space are given, as well as the development of a method for identifying trolling accounts in social networks is described and tested in the Internet space.

Key words: digital economy, digitalization, virtual world, virtual aggression, cyber aggression, trolling, digital trolling index.

Экспоненциальное развитие технологий и появление глобальной сети Интернет во второй половине XX века позволило человечеству перейти на новый высококачественный уровень коммуникаций, опосредованный определенными особенностями онлайн-поведения. Такие особенности поведения способствуют появлению и развитию негативных последствий коммуникации, как виртуальная

агрессия и троллинг

Троллинг с точки зрения научного обоснования рассматривается как провокационное сообщение, зачастую, с элементами оскорбления в адрес конкретного человека в социальном Интернет–пространстве.

По данным исследований, «троллинг» занимает третье место в списке причин смерти подростков в Соединенных Штатах Америки на сегодняшний день, а в Российской Федерации около 70% детей в возрасте от 12 до 16 лет становились жертвами или сами травили сверстников в социальной сети. В других странах ситуация приближается к таким же показателям.

Исходя из вышеизложенного, данная проблема имеет глобальный характер, а ее последствия могут быть необратимыми.

Феномен троллинга тесно связан с проблемой анонимности в Интернете: с одной стороны, анонимность является одной из ведущих ценностей Интернет–среды, с другой, интернет–мошенники, в том числе занимающиеся троллингом, активно пользуются этой ценностью в своих корыстных целях. С данной точки зрения вышеупомянутая проблема двойственной анонимности в глобальной сети становится проблемой, угрожающими психическому здоровью человека и является социально-значимой задачей для современного мира, что определяет актуальность данной работы [8].

Объектом исследования является феномен интернет-троллинга.

Предметом исследования являются методики выявления троллинговых аккаунтов, интернет-троллей и потенциальных жертв троллинга.

Впервые в научной литературе использование слова «троллинг» произошло в 1996 году и принадлежит американскому специалисту по современным медиа-средствам Джудит Донат [32], которая в своих исследованиях описывала некоторые нестандартные ситуации коммуникации людей в среде общения Usenet. Донат дала описание явлению троллинга как умышленно вредоносную ложь, отмечая, что «тролли» способствуют быстрому снижению доверия и терпимости к людям, а также способствуют развитию паранойи в онлайн сообществе [2].

В настоящее время термин «троллинг» определяется как форма социальной провокации или издевательства в сетевом пространстве, использующаяся как персонифицированными участниками, заинтересованными в большей узнаваемости, публичности, эпатаже, так и анонимными пользователями без возможности их идентификации. По мнению психолога, изучающего кибер-агрессию, Внебрачных Р.А., троллинг можно рассматривать

как социальное подстрекательство и психологическое насилие в сети Интернет [5].

Основными местами существования явления троллинга и его модификаций выступают различные тематические форумы, конференции, социальные сети, порталы, чаты и новостные сайты.

Человек–троль пытается представить себя типичным пользователем в социальных, который разделяет общие интересы и проблемы группы либо сообщества. Тролли могут нанести существенный вред коммуникации во многих направлениях: испортить обсуждение, распространить вредный совет либо деструктивную идею, разрушить чувство взаимного доверия в сообществе.

В настоящее время явление троллинга и распространение его модификаций становится проблемой, угрожающей психическому, физическому и эмоциональному здоровью пользователей сети Интернет, с широким спектром последствий от избегания реального общения до возникновения депрессивных состояний и суицидальных наклонностей.

Самой уязвимой возрастной группой попадания под воздействие троллинга считается подростковый возраст от 12 до 16 лет [6]. По утверждениям психологов, в нашей стране от явления троллинга страдает каждый третий ребенок [8], а согласно статистике государственной организации «Роскачество» на декабрь 2019 года около 70% российских подростков становились жертвами или сами травили сверстников в социальной сети [9] и 42% находились в непосредственном визуальном контакте друг с другом в реальной жизни [1].

Поведение детей, попавших под влияние троллинга обладает также определенными особенностями: они становятся уязвимыми, замкнутыми, агрессивными и скрытными, в их образе жизни происходят резкие изменения: они начинают менять свою внешность, стиль одежды, их круг общения сужается до виртуального, ребенок закрывается в собственной комнате, большее количество времени, в том числе, на уроках в школе, проводит в Сети, не выходит на контакт с родителями и преподавателями, не реагирует на замечания.

Суть данного развлечения заключается в том, что дети выполняют определенные задания куратора переходя на «своеобразные» уровни. Финальный уровень является фатальным – совершение самоубийства. В роли кураторов игры «Синий кит» чаще всего выступают дети–подростки, такие же, как и их гипотетические жертвы. Изначально кураторы приглашают детей в глобальных сообществах сети поиграть в интересный онлайн–квест, далее, создают специальные паблики и конференции для коммуникации с участниками

игры, общаются в личных сообщениях, оставляют комментарии и записи на странице потенциальной жертвы.

Согласно данным Центра спасения детей от кибер-преступлений 58% детей в возрасте от 12 до 16 лет скрывают от родителей свои действия в сети [7, 9], предпринимая для этого определенные усилия.

Исходя из вышеизложенного, с целью выявления суицидальных наклонностей у ребенка, психологических отклонений, предотвращения агрессивного поведения детей внутри Сети, негативного влияния на сверстников и предотвращения фатальных исходов педагоги школ вынуждены вести мониторинг социальных сетей детей, т.к. при совершении определенных поступков подростками их родители винят в случившемся школу, учителей и прежде всего, классного руководителя ребенка.

У детей, подвергавшихся психологическому влиянию со стороны троллей на страницах обнаружены записи, связанные с группой смерти, цитаты о загробном мире, смысле жизни, одиночестве, комментарии людей со странными фамилиями и аватарками, фотографии порезов рук, рисунки бабочек и китов, информация о конфликтах со сверстниками, родителями.

Подростки указывают о себе ложную информацию на странице (меняют фамилии и имена, фото профиля, город проживания, дату рождения и т.д.), удаляют из друзей знакомых людей, оставляя лишь малое количество виртуальных, выходят из социальных сообществ и многое другое.

За 2019 год преподавателями школ России было установлено 928 несовершеннолетних участников групп и чатов, в отношении которых совершались действия противоправной и преступной направленности, имеющих страницы в социальных сетях, участников сообществ в каналах различных мессенджеров, 3582 сообщества, где идет активное склонение детей к самоубийству, в том числе публикациями шок контента, размещением информации специализированного типа склоняющей, пропагандирующей, романтизирующей суицид, одиночество, отрешение от мира, как единственный метод разрешения трудных жизненных ситуаций [5].

В настоящее время существующих методов для борьбы с явлением троллинга недостаточно, т.к. они основаны на работе администраторов сообществ и научных сотрудников, поэтому, в рамках данной работы мы разработали собственную методику выявления троллей в сети на основе нечеткой логики.

Методика оценки цифрового индекса троллинга основывается на поведении людей, находящихся в группе риска и наиболее подходит для анализа страниц подростков.

На первом шаге преподаватель школы совместно с детьми на классном часе проходят тест «DigiFriends», разработанный Высшей школой экономики для социальной сети ВКонтакте [5]. В основе данного теста лежит анкета, вопросы которой касаются онлайн-поведения школьника.

Вторым шагом данной методики является автоматический сбор открытых данных о школьниках с их социальных профилей с помощью специального программного обеспечения или онлайн-сервиса.

Третий шаг методики – визуализация взаимодействия классного коллектива в социальном пространстве.

Четвертым шагом нашей методики является анализ социальных страниц детей, выделенных на третьем шаге и экспертная оценка их поведения в сети и в классе на основании определенных принципов. Выбранные критерии не имеют четких разграничений, поэтому, за основу была взята нечеткая логика.

Разработанная модель и примеры апробации методики на основе нечеткой логики представлены на рисунках 1 и 2.

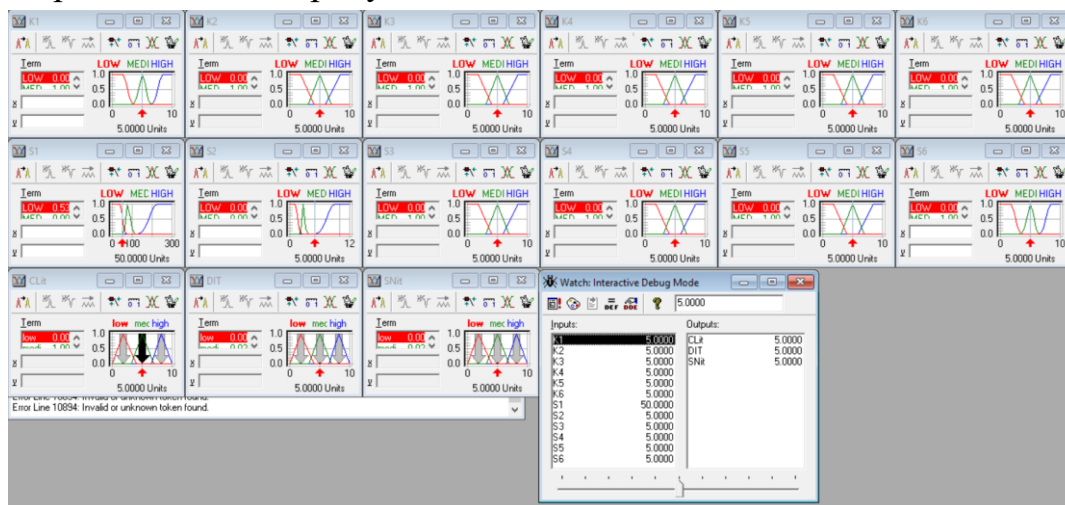


Рисунок 1 – Разработанная модель DIT

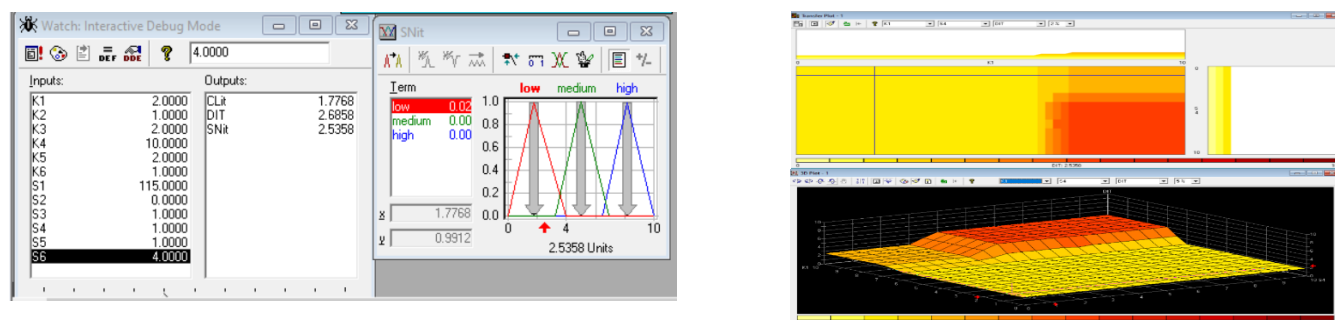


Рисунок 2 – Пример апробации методики

Согласно выставленным экспертным оценкам наша модель выдала результаты промежуточных и итоговой переменной. Значение промежуточной переменной Clit равно 1,7768, что обозначает хорошее поведение в классе. Значение промежуточной переменной Snit, отвечающей за профиль ребенка, равно 2,5358. Социальная страница не вызывает подозрения у преподавателей. Единственным недостатком достоверности аккаунта является неверно указанный возраст (15 лет). Значение цифрового индекса троллинга равно 2,6858.

Открытие данных графиков визуализации переменных позволяет анализировать характеристики входов и выходов нечеткой системы путем отображения поверхностей в пространстве управления. Две входные переменные откладываются по оси абсцисс и оси ординат соответственно, а значение выбранной выходной переменной представляется в виде раскрашенной поверхности, при этом, цветовая гамма соответствует степени взаимосвязи входных и выходных переменных. Графики отражают комплексную картину зависимости выходных переменных от входных и являются удобным инструментом анализа. Данные окна также могут служить способом визуализации оценки влияния изменения значений входных нечетких переменных на значение одной из выходных нечетких переменных.

Данная методика прошла успешную апробацию в реальной жизни и применяется в двух образовательных организациях Курганской области.

Разработанная модель DIT позволяет выявить пользователей, попавших под явление троллинга в сети–Интернет и своевременно оказать им психологическую поддержку, а также, определить агрессивные троллинговые и недостоверные аккаунты в социальных сетях, используя экспертную оценку социального профиля в Интернет-пространстве без идентификации реального поведения человека в жизни.

В дальнейшем данная методика может быть внедрена в деятельность не только общеобразовательных организаций, но и в Интернациональный центр спасения детей от кибер-преступлений, а также, в онлайн–жизнь обычных равнодушных пользователей социальных сетей к явлению троллинга.

Библиографический список

1. Fuzzy Logic: Четкие решения нечеткой логики [Электронный ресурс]. – Режим доступа: http://www.bacnet.ru/knowledgebase/articles/index.php?ELEMENT_ID=653, свободный.

2. Judith S. Donath, MIT Media Lab. Identity and Deception in the Virtual Community. Prepared for: Kollock, P. and Smith M. (eds). Communities in Cyberspace. London: Routledge, 1996, 11(2):246–278.

3. Zelenkauskaitė A, Niezgodą B. Stop kremlin trolls: Ideological trolling as calling out, rebuttal, and reactions on online news portal commenting. First Monday, 2017, 22(5).

4. Внебрачных Р. А. Троллинг как форма социальной агрессии в виртуальных сообществах [Текст] / Р.А. Внебрачных // Вестник Удмуртского университета. Философия. Социология. Психология. Педагогика. — Ижевск: Удмуртский государственный университет, 2012. — 1. — С. 48—51.

5. Исследовательский проект Высшей школы экономики Digital Friends. [Электронный ресурс]. – Режим доступа: <https://linis.hse.ru/soft-linis/> – диджифрендс, свободный.

6. Конышева Л.К., Назаров Д.М. Основы теории нечетких множеств: Учебное пособие. -Изд. Питер, 2011. - 190 с.

7. Мелешкина И.И. Корреляция киберпространства и современного поколения России. Международная научно-практическая конференция: Внедрение результатов инновационных разработок: проблемы и перспективы. 2019.

8. Мелешкина И.И. Троллинг и этнотроллинг как формы кибер-агрессии. II Международная научно-практическая конференция «Актуальные направления научных исследований». 19 мая 2018.

9. Мелешкина И.И., Бегичева С.В. Кибербуллинг и астротурфинг как виртуальная агрессия. I научно-практическая конференция. РАНХиГС. Сборник: Современные информационные технологии: проблемы и перспективы развития. 2017.